

Basit Doğrusal Regresyon Analizi

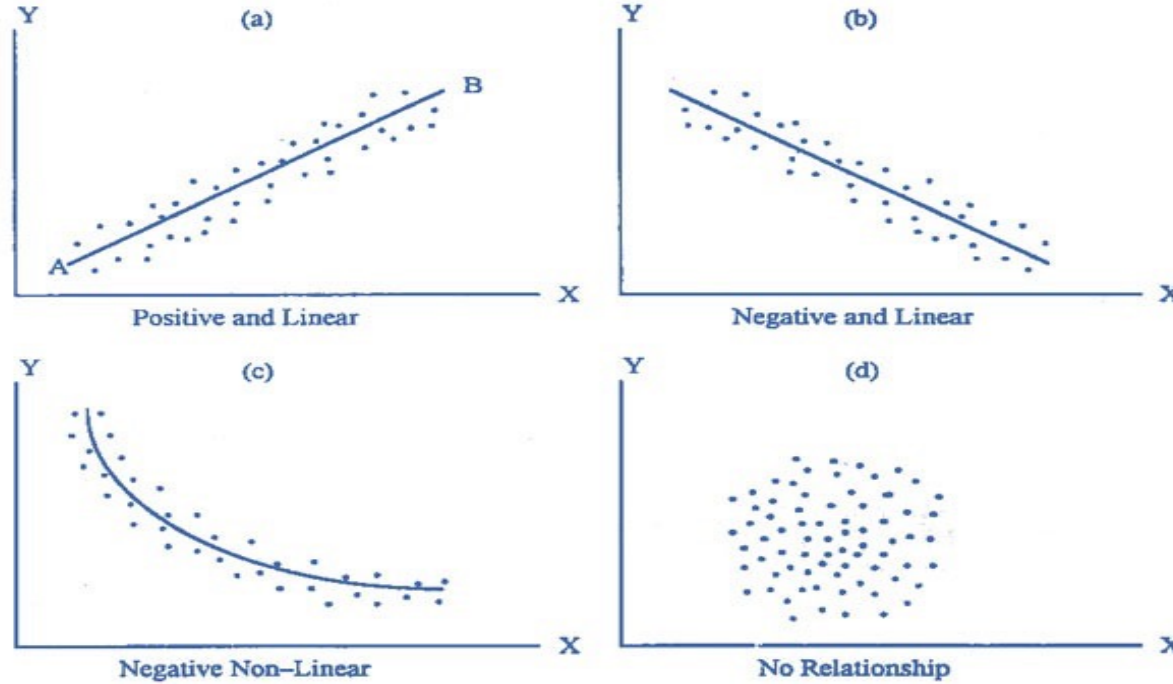
- Doğrusal regresyon, korelasyondan sonraki adımdır.
- Başka bir değişkenin değerine bağlı olarak bir değişkenin değerini tahmin etmek istediğimizde kullanılır.
- Tahmin etmek istediğimiz değişken bağımlı değişken (ya da bazen sonuç değişkeni) olarak adlandırılır.
- Diğer değişkenin değerini tahmin etmek için kullandığımız değişkene bağımsız değişken (veya bazen tahmin edici değişken) denir.
- Örneğin, sınav performansının gözden geçirme süresine bağlı olarak tahmin edilip edilemeyeceğini; sigara tüketiminin sigara içme süresine bağlı olarak tahmin edilip edilemeyeceğini vb. anlamak için doğrusal regresyon kullanabilirsiniz.
- Tek bir bağımsız değişken yerine iki veya daha fazla bağımsız değişkeniniz varsa, çoklu regresyon kullanmanız gerekir.

Varsayımları

- Verilerinizi doğrusal regresyon kullanarak analiz etmeyi seçtiğinizde, sürecin bir parçası, analiz etmek istediğiniz verilerin gerçekten doğrusal regresyon kullanılarak analiz edilebileceğinden emin olmak için kontrol etmeyi içerir.
- Bunu yapmanız gerekir çünkü doğrusal regresyonun geçerli bir sonuç vermesi için gerekli olan yedi varsayımı verileriniz "geçerse" doğrusal regresyonu kullanmak uygun olur.

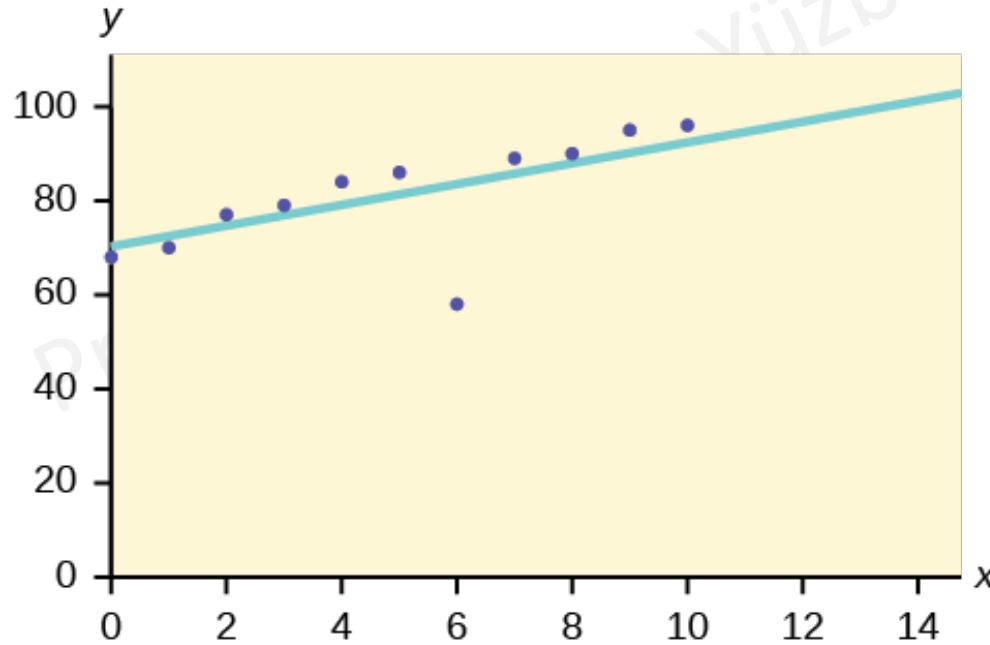
- **Varsayım #1: Bağımlı değişken**iniz **sürekli** düzeyde ölçülmelidir (yani ya bir **aralık** ya da **oran** değişkenidir). **Sürekli değişkenlere** örnek olarak revizyon süresi (saat olarak ölçülür), zeka (IQ puanı kullanılarak ölçülür), sınav performansı (0'dan 100'e kadar ölçülür), ağırlık (kg olarak ölçülür) ve benzerleri verilebilir.
- **Varsayım #2: Bağımsız değişken**iniz aynı zamanda **sürekli** düzeyde ölçülmelidir (yani ya bir **aralık** ya da **oran** değişkenidir)

- **Varsayım #3:** İki deęişken arasında **doęrusal** bir **ilişki** olması gerekir. İki deęişkeniniz arasında doęrusal bir ilişki olup olmadığını kontrol etmenin çeşitli yolları olsa da, SPSS Statistics kullanarak baęımlı deęişkeni baęımsız deęişkeninize karşı çizebileceğiniz bir dağılım grafięi oluşturmanızı ve ardından doęrusallığı kontrol etmek için dağılım grafięini görsel olarak incelemenizi öneririz. Dağılım grafięiniz aşağıdakilerden birine benzevebilir:



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)

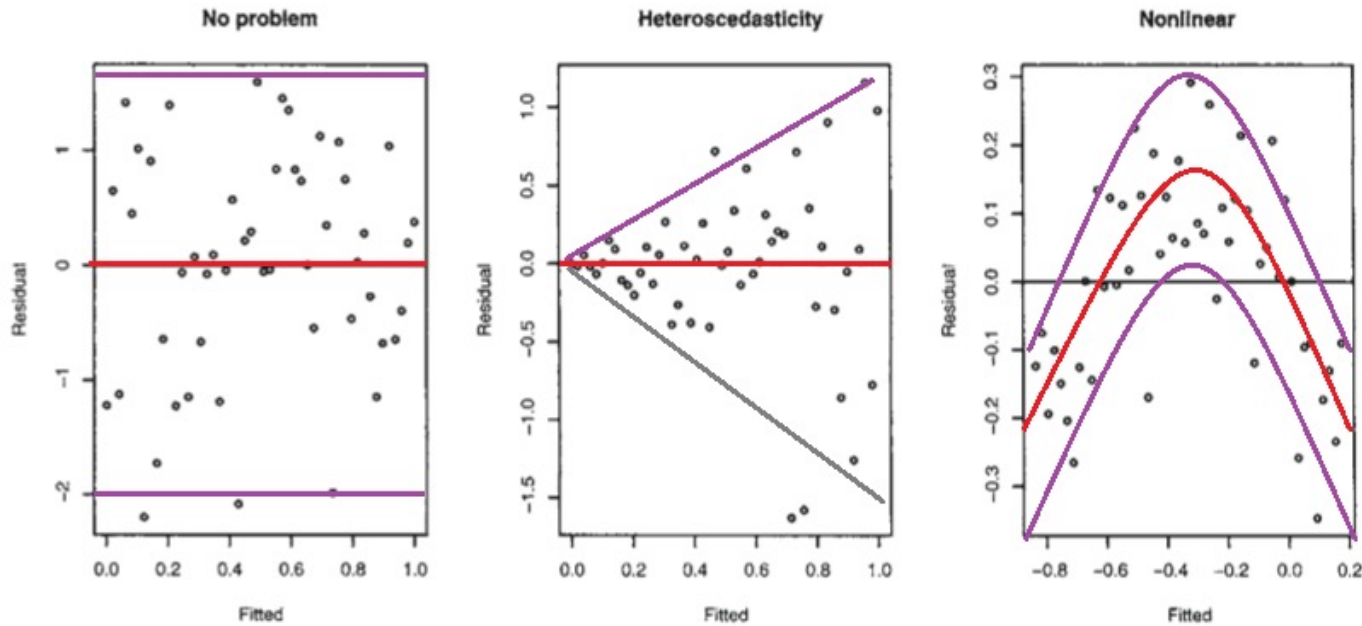
- **Varsayım #4: Önemli aykırı değerler olmamalıdır.** Aykırı değer, regresyon denkleminin öngördüğü değerden çok farklı bir bağımlı değişken değerine sahip gözlenen bir veri noktasıdır. Bu nedenle, bir aykırı değer, dağılım grafiğinde regresyon doğrusundan (dikey olarak) uzakta olan ve aşağıda vurgulandığı gibi büyük bir kalıntıya sahip olduğunu gösteren bir nokta olacaktır:



[This Photo](#) by Unknown Author is licensed under [CC BY](#)

- **Varsayım #5:** Gözlemlerin bağımsızlığına sahip olmalısınız, bunu SPSS Statistics kullanarak çalıştırabileceğiniz basit bir test olan Durbin-Watson istatistiğini kullanarak kolayca kontrol edebilirsiniz. Durbin-Watson istatistiğinin sonucunun nasıl yorumlanacağını gelişmiş doğrusal regresyon kılavuzumuzda açıklıyoruz.
- **Varsayım #6:** Regresyon çizgisinin artıklarının (hatalarının) yaklaşık olarak normal dağıldığını kontrol etmeniz gerekir (bu terimleri gelişmiş doğrusal regresyon kılavuzumuzda açıklıyoruz). Bu varsayımı kontrol etmek için kullanılan iki yaygın yöntem arasında bir histogram (normal eğri eklenmiş) ya da Normal P-P Çizimi kullanmak yer alır. Yine, geliştirilmiş doğrusal regresyon kılavuzumuzda (a) ister histogram (üst üste bindirilmiş normal eğri ile) ister Normal P-P Plot kullanın, SPSS Statistics kullanarak bu varsayımı nasıl kontrol edeceğinizi göstereceğiz; (b) bu diyagramları nasıl yorumlayacağınızı açıklayacağız; ve (c) verileriniz bu varsayımı karşılamıyorsa olası bir çözüm sunacağız.

- **Varsayım #7:** Verilerinizin homoscedasticity göstermesi gerekir; bu, en iyi uyum çizgisi boyunca varyansların çizgi boyunca ilerledikçe benzer kaldığı durumdur. Bunun ne anlama geldiği ve verilerinizin eş varyanslılığının nasıl değerlendirileceği hakkında daha fazla bilgiyi gelişmiş doğrusal regresyon kılavuzumuzda açıklarken, üç basit örnek sunan aşağıdaki üç dağılım grafiğine bir göz atın: ikisi varsayımı karşılamayan (heteroscedasticity olarak adlandırılır) ve biri bu varsayımı karşılayan (homoscedasticity olarak adlandırılır) veriler:



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)

Çoklu Regresyon

- Çoklu regresyon, basit doğrusal regresyonun bir uzantısıdır. İki veya daha fazla başka değişkenin değerine bağlı olarak bir değişkenin değerini tahmin etmek istediğimizde kullanılır. Tahmin etmek istediğimiz değişken bağımlı değişken (veya bazen sonuç, hedef veya kriter değişkeni) olarak adlandırılır. Bağımlı değişkenin değerini tahmin etmek için kullandığımız değişkenlere bağımsız değişkenler (veya bazen tahmin edici, açıklayıcı veya regresör değişkenler) denir.

- Örneğin, sınav performansının gözden geçirme süresi, sınav kaygısı, derse katılım ve cinsiyete bağlı olarak tahmin edilip edilemeyeceğini anlamak için çoklu regresyon kullanabilirsiniz.
- Alternatif olarak, günlük sigara tüketiminin sigara içme süresi, sigaraya başlama yaşı, içici tipi, gelir ve cinsiyete göre tahmin edilip edilemeyeceğini anlamak için çoklu regresyon kullanabilirsiniz.
- Çoklu regresyon ayrıca modelin genel uyumunu (açıklanan varyans) ve tahmin edicilerin her birinin açıklanan toplam varyansa göreceli katkısını belirlemenizi sağlar.
- Örneğin, sınav performansındaki varyasyonun ne kadarının tekrar süresi, sınav kaygısı, derse katılım ve cinsiyet tarafından "bir bütün olarak" açıklanabileceğini ve aynı zamanda her bir bağımsız değişkenin varyansı açıklamadaki "göreceli katkısını" bilmek isteyebilirsiniz.

Varsayımlar

- Varsayım #1: Bağımlı değişkeniniz sürekli bir ölçekte ölçülmelidir
- **Varsayım #2: Sürekli** (yani **aralık** veya **oran** değişkeni) veya kategorik (yani **sıra** veya **nominal** değişken) olabilen **iki veya daha fazla bağımsız** değişkeniniz vardır. **Sürekli** ve **sıralı** değişken örnekleri için yukarıdaki maddeye bakınız. **Nominal değişkenlere** örnek olarak cinsiyet (örn. 2 grup: erkek ve kadın), etnik köken (örn. 3 grup: Beyaz, Afro-Amerikan ve Hispanik), fiziksel aktivite düzeyi (örn. 4 grup: hareketsiz, düşük, orta ve yüksek), meslek (örn. 5 grup: cerrah, doktor, hemşire, diş hekimi, terapist) vb. verilebilir.

- **Varsayım #3: Gözlemlerin bağımsızlığına** (yani **artıkların bağımsızlığına**) sahip olmalısınız; bunu SPSS Statistics kullanarak çalıştırabileceğiniz basit bir test olan Durbin-Watson istatistiğini kullanarak kolayca kontrol edebilirsiniz.
- **Varsayım #4:** (a) bağımlı değişken ile bağımsız değişkenlerinizin **her biri** arasında ve (b) bağımlı değişken ile bağımsız değişkenler arasında **doğrusal** bir **ilişki** olmalıdır. Bu doğrusal ilişkileri kontrol etmenin birçok yolu olmakla birlikte, SPSS Statistics kullanarak **dağılım graf**ikleri ve kısmi **regresyon** grafikleri oluşturmanızı ve ardından doğrusallığı kontrol etmek için bu dağılım grafiklerini ve kısmi regresyon grafiklerini görsel olarak incelemenizi öneririz. Dağılım ve kısmi regresyon grafiklerinizde görüntülenen ilişki doğrusal değilse, ya doğrusal olmayan bir regresyon analizi yapmanız ya da verilerinizi "dönüştürmeniz" gerekecektir ki bunu SPSS Statistics kullanarak yapabilirsiniz.
- **Varsayım #5:** Verilerinizin **homoscedasticity** göstermesi gerekir; bu, en iyi uyum çizgisi boyunca varyansların çizgi boyunca ilerledikçe benzer kaldığı durumdur.
- **Varsayım #6:** Verileriniz, birbiriyle yüksek korelasyona sahip iki veya daha fazla bağımsız değişkeniniz olduğunda ortaya çıkan **çoklu doğrusallık** göstermemelidir. Bu durum, bağımlı değişkende açıklanan varyansa hangi bağımsız değişkenin katkıda bulunduğu anlaşılmasında sorunlara ve çoklu regresyon modelinin hesaplanmasında teknik sorunlara yol açar.

- **Varsayım #7: Önemli aykırı değerler, yüksek kaldıraç noktaları veya yüksek etkili noktalar** olmamalıdır. Aykırı değerler, kaldıraç ve etkili noktalar, çoklu regresyon analizi yapmak istediğinizde veri setinizde bir şekilde olağandışı olan gözlemleri temsil etmek için kullanılan farklı terimlerdir. Olağandışı **noktaların** bu farklı sınıflandırmaları, regresyon çizgisi üzerinde sahip oldukları farklı etkileri yansıtır.
- **Varsayım #8:** Son olarak, artıkların (hataların) yaklaşık olarak normal dağıldığını kontrol etmeniz gerekir. Bu varsayımı kontrol etmek için kullanılan iki yaygın yöntem şunlardır: (a) bir histogram (üst üste bindirilmiş normal eğri ile) ve bir Normal P-P Çizimi; veya (b) studentized artıkların bir Normal Q-Q Çizimi.